Let us consider the conditions under which Type I Deterrence might fail.

Against an irrational opponent, of course, a deterrent policy might have no effect whatever. Let us consider, rather, the failure to deter--by the threat of a retaliatory or pre-emptive nuclear strike--a rational opponent from attempting to strike first.

To simplify the problem, let us assume that the opponent--the SU--has only two alternative strategies, which we will label: Strike, and Wait. "Strike" means to launch an all-out nuclear attack upon the US, in an attempt to strike first. "Wait" comprises every other strategy, i.e., every SU strategy which does not involve a first strike against the US.

In comparing these two strategies, let us imagine that the SU considers only two contingencies affecting their outcomes: the concurrent choice by the US either to Wait or to Strike. In general, we will assume that neither "player" can predict his opponent's choice with certainty, and that neither believes that his opponent can do so. Neither, in other words, "moves first." The Soviet's problem can be represented as the choice of a column in the following matrix:

		SU: Wait	SU: Strike
US	Wait	v_{11}	v_{21}
	Strike	v_{12}	v_{22}

The "v" elements in the matrix represent Soviet "payoffs" (US payoffs are not shown; they are not merely the negative of Soviet payoffs, i.e., the game is not zero-sum). We will assume that these payoffs are numbers with the properties of von Neumann-Morgenstern utilities. That means, for example, that v_{21} will equal $\frac{1}{2}(v_{11} + v_{12})$ if and only if the SU is indifferent between the outcome v_{21}

(I.e., they incorporate information about the choices the SU would actually make among "lottery tickets" which offer the various outcomes as "prizes" with different probabilities). Furthermore, we will assume that the SU distinguishes subjectively between the "likelihoods" of the possible US choices, that it attaches different "degrees of likelihood" to the US strategies, and that these degrees of likelihood can be represented by "subjective probabilities": i.e., numbers having the properties of probabilities, to be measured or estimated (like the utilities) by Soviet choices among "lottery tickets", in the manner indicated by L.J. Savage.

These are restrictive assumptions, which definitely oversimplify the problem; yet, as we shall see, the analysis can become rather complicated even so. In general, the problem of estimating the values of the few variables considered here could be very imposing. However, in the general discussion of deterrence, judgments are frequently introduced which amount to "estimates" of these very factors; an immediate application of this simple model, then, may be to help in deducing significant implications of such judgments.

Under these assumptions as to data, it follows that the SU will "maximize its expected utility payoff." (The utility numbers and probability numbers have been chosen, let us remember, precisely to make this an accurate description of actual Soviet choice behavior under uncertainty). I.e., it will choose one action over another when and only when the sum of the payoffs associated with the first, weighted by their probabilities, exceeds the weighted utility sum for the second.

Thus, the SU will choose Strike rather than Wait in this situation if and only if: U(Strike) > U(Wait), which will be the case if and only if:

$$(1-q)\cdot V_{21} + q\cdot V_{22} > (1-q)\cdot V_{11} + q\cdot V_{12}$$

where q is the subjective probability the SU assigns to a US choice of Strike.

Whether this condition might ever hold would depend on the possible range of values for the payoffs and for q. For example, the opinion is sometimes expressed that a major nuclear war--no matter which side started it, or under what conditions of surprise--would result in "the end of all life on earth." If we fix the scale and origin of the SU utility index by assigning arbitrarily the value 0 to the "status quo" outcome (the result of a mutual choice of Wait) and -100 to the "worst conceivable" outcome, the above technological assumption would imply a payoff matrix similar to the following one:

		SU: Wait	SU: Strike
US	Wait	0	-100
	Strike	-100	-100

The implications for Soviet rational choice may be obvious, but let us spell them out for practice in working with the model. First, we observe that the condition:

$$(1-q)\cdot(-100) + q\cdot(-100) > (1-q)\cdot(0) + q\cdot(-100)$$

can never hold, for any value of q between 0 and 1. Hence, ~~Wait must always be preferred to Strike~~. Strike can never be preferred to Wait.

This might be deduced directly from the matrix if we observe that the strategy Wait "dominates" Strike. Under every given contingency (i.e., US choice) the outcome of Wait under that contingency is preferable to that of Strike; hence, the likelihoods of the various contingencies cannot affect the SU choice. ~~With these payoffs, it would be irrational for the SU to choose Strike~~

The payoffs imply that there is no advantage to the SU in a "pre-emptive" or "preventive" strike; even if they were certain that the US would choose Strike, the SU would have no incentive to choose Strike. (This does not mean that the US would be untouched; the policy represented by Wait precludes, by definition, a first strike, but it would generally include provision for a second strike: i.e., a retaliatory blow with remaining forces, if struck).

Under these conditions it would be strictly irrational for the SU to choose Strike. This is, of course, the conclusion of those who believe in these technological payoffs; and their conclusion is consistent with their assumption.

However, there is considerable question as to whether these technological predictions are sound. There is considerable reason to believe that the physical outcome of striking first (and su

Let us suppose, however, that these technological predictions are unsound. Suppose that the physical consequences to the attacker of striking first (and suffering his opponent's retaliatory strike with remaining forces) look very different than the consequences of choosing Wait and being struck with the full weight of his opponent's first strike. To take civilian deaths as an index of physical damage, let us suppose that the outcome of choosing Strike when the US chose Wait might be 40 million SU dead, compared to 200 million dead if the US chose Strike and the SU chose Wait (neither situation implying, we are assuming, the "end of all life on earth").

Physical outcomes still do not determine utility payoffs; and it is the latter which are relevant to the SU decision. It might be argued (estimated) that the Soviets would might still attach the same utility number to these different outcomes--if not -100, something close to it--so that Wait would still dominate Strike. The SU, this would imply, would still see no advantage in choosing Strike even if they were certain that the US would strike. But such an assumption does not seem compelling. Let us assume instead that such a large discrepancy in physical outcome would be reflected in the utility payoffs corresponding to these outcomes. The utilities might then appear:

	Wait	Strike	W	S
Wait	0	-200	0	-70
Strike	-70	-90	-100	-90

1/8

To arrive at any such outcomes as these, one must make specific assumptions about the nature of the different strategies. Let us assume that Strike implies some sort of high-alert posture as well as a decision to attack, and that Wait implies a lower alert. The outcome of choosing Strike when the opponent has also chosen Strike will depend on the precise nature of the offensive and defensive systems, including the reliance on warning and alert; it will also reflect the timing of the two attacks, which may depend on more or less random factors beyond the control of the top decision-makers. However, we can assume that the outcome will not be worse than that of choosing Wait when the opponent chooses Strike, nor better than that of choosing Strike against the opponent's Wait; i.e., it will lie between these payoffs, as shown in this example.

In this case, Wait no longer dominates Strike. Neither strategy is dominant. Which strategy will appear preferable to the Soviets will now depend upon q, the Soviet expectation of a US choice of Strike.